AFRL-RI-RS-TR-2014-045

ROBUST ROBOT CONTROL USING MULTIPLE MODEL-BASED POLICY OPTIMIZATION AND FAST VALUE FUNCTION-BASED PLANNING

CARNEGIE MELLON UNIVERSITY

MARCH 2014

FINAL TECHNICAL REPORT

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED

STINFO COPY

AIR FORCE RESEARCH LABORATORY
INFORMATION DIRECTORATE

■ **AIR FORCE MATERIEL COMMAND** ■ **UNITED STATES AIR FORCE** ■ **ROME, NY 13441**

NOTICE AND SIGNATURE PAGE

AFRL-RI-RS-TR-2014-045 HAS BEEN REVIEWED AND IS APPROVED FOR PUBLICATION IN ACCORDANCE WITH ASSIGNED DISTRIBUTION STATEMENT.

FOR THE DIRECTOR:

/ S /
PETER J. ROCCI
Work Unit Manager

/ S /
MICHAEL J. WESSING
Deputy Chief, Information Intelligence Systems & Analysis Division
Information Directorate

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	**2. REPORT TYPE**	**3. DATES COVERED** *(From - To)*
MARCH 2014	FINAL TECHNICAL REPORT	SEP 2012 – SEP 2013

4. TITLE AND SUBTITLE

ROBUST ROBOT CONTROL USING MULTIPLE MODEL-BASED POLICY OPTIMIZATION AND FAST VALUE FUNCTION-BASED PLANNING

5a. CONTRACT NUMBER FA8750-12-1-0307

5b. GRANT NUMBER N/A

5c. PROGRAM ELEMENT NUMBER 62702E

6. AUTHOR(S)

Christopher Atkeson

5d. PROJECT NUMBER ROBO

5e. TASK NUMBER PR

5f. WORK UNIT NUMBER 09

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Carnegie Mellon University
500 Forbes Ave
Pittsburgh, PA 15213

8. PERFORMING ORGANIZATION REPORT NUMBER
N/A

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Air Force Research Laboratory/RIED
525 Brooks Road
Rome NY 13441-4505

10. SPONSOR/MONITOR'S ACRONYM(S)
AFRL/RI

11. SPONSORING/MONITORING AGENCY REPORT NUMBER
AFRL-RI-RS-TR-2014-045

12. DISTRIBUTION AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited. This report is the result of contracted fundamental research deemed exempt from public affairs security and policy review in accordance with SAF/AQR memorandum dated 10 Dec 08 and AFRL/CA policy clarification memorandum dated 16 Jan 09.

13. SUPPLEMENTARY NOTES

14. ABSTRACT
This report describes the research findings of Team Steel, the group led by PI Christopher G. Atkeson in the DARPA Virtual Robotics Challenge (VRC). They developed human-like walking and robust walking on rough terrain, and automated driving in the VRC context. They developed a rough terrain footstep planner, a decoupled approach to state estimation, and an optimization based real-time walking controller for a full size 3D humanoid robot. They showed that optimal stepping trajectories and trajectory cost for a walking biped robot on rough terrain can be encoded as simple quadratic functions of initial state and footstep sequence.

15. SUBJECT TERMS
Robot Control, Humanoid Robot, Robot Walking, Robot Driving, Footstep Planning, Robot State Estimation

16. SECURITY CLASSIFICATION OF: a. REPORT	b. ABSTRACT	c. THIS PAGE	**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON** / **19b. TELEPHONE NUMBER** *(Include area code)*
U	U	U	UU	12	**PETER J. ROCCI** / **315-330-4654**

Table of Contents

1.0 Summary

This report describes the research findings of Team Steel, the group led by PI Christopher G. Atkeson in the DARPA Virtual Robotics Challenge (VRC). We made excellent progress in developing human-like walking and robust walking on rough terrain, and automated driving in the VRC context. We developed a rough terrain footstep planner, a decoupled approach to state estimation, and an optimization based real-time walking controller for a full size 3D humanoid robot. We showed that optimal stepping trajectories and trajectory cost for a walking biped robot on rough terrain can be encoded as simple quadratic functions of initial state and footstep sequence. Our paper on our walking algorithm won "Best Oral Paper Award" at Humanoids 2013. We are applying our ideas to manipulation, walking, and climbing a ladder in the DARPA Robotics Challenge (DRC) through our participation in the Worcester Polytechnic Insittute WRECS Team.

2.0 Introduction

We participated in the DARPA Virtual Robotics Challenge. This challenge consisted of several simulated robot tasks. We focused on the rough terrain task and the driving task. The rough terrain task consisted of simulated walking across mud, hilly terrain, and among obstacles to reach a position target. The driving task consisted of getting into the vehicle, driving along a road, getting out of a vehicle at a target, and walking through a gate. The robot used was a simulation of BDI's Atlas.

3.0 Methods, Assumptions, and Procedures

Our methods and procedures involved programming in C++ for the Gazebo simulator, within the published constraints of the DARPA Virtual Robotics Challenge.

We built on our previous work on humanoid robot control, and our software is an evolution of an existing design developed for the DARPA Maximum Mobility and Manipulation (M3) Program. Our high-level software architecture includes at the highest level a finite state machine to supervise task execution and handle system level errors that may require a reset of all or large portions of the system. The second level includes a finite state machine to determine what task we are performing, and the third level provides task specific finite state machines that implement the phases of each particular task. Below this level the architecture is determined by the nature of the task. Error detection and handling is done at all levels. The third level uses task state transitions to handle errors within a task. The second level detects failures in task execution such as stuck states, terminates execution of failed tasks, and initiates execution of recovery tasks. The top level acts like a watchdog interrupt for the entire system, detecting failures of the entire control system, such as when second level task monitors are failing.

In terms of obstacle avoidance and path and motion planning, we built on our previous work in both the DARPA M3 Program and the DARPA Learning Locomotion Program. Motion planning at the most detailed level is too slow to consider a wide range of alternative plans. For the DARPA M3 Program we developed a decoupled dynamic programming (DP) approach for legged locomotion. Given a terrain cost map, the DP approach can globally optimize where and when to put the swing foot down on the next step by optimizing the choice of action using a “value function” or “cost-to-go” that was calculated by dynamic programming offline in advance of deploying the robot. The use of the precomputed value function takes less than a millisecond, a tiny fraction of the 100s of milliseconds it takes to actually swing the foot.

We extended our previous quadratic programming-based approach to walking. At the core inverse dynamics is done, while obeying friction cone and foot tipping (Center of Pressure (COP)/Zero Moment Point (ZMP)) constraints. We focused on center of mass (COM) control, using center of mass acceleration to determine necessary foot forces. In cases with support, the contact forces are allocated to maximize the worst case distance to each contact friction cone. Center of mass motion is planned and optimized using simple dynamics models, such as the Linear Inverted Pendulum (LIPM) model, the LIPM model augmented with vertical movement, and the LIPM model augmented with angular momentum.

4.0 Results and Discussion

We developed a footstep planner that worked well on the VRC rough terrain task. Two key ideas were developed. The first is to estimate a cost that approximated the cost an optimizer would generate to cross a terrain. We used simple functions modeled on the metabolic cost for biological systems. Amazingly enough, this worked well, even though one would imagine that metabolic cost was not important for transient behavior. We also developed cost functions for terrain that estimated the risk of falling, for example to what extent will the terrain support the foot only on a point or edge contact, and allow the foot to rock. Details are presented in [2].

We developed a decoupled approach to state estimation that allows us to efficiently use information from the full body dynamics. We trade partial information loss for computational efficiency. The main idea is to decouple the full body state vector into several independent state vectors, such as the center of mass dynamics and the individual joint dynamics. Dynamic coupling between the joints is included in the state estimators for the joints. The full body dynamics are projected into a suitable subspace that depends on the contact state. Each decoupled state vector can be estimated very efficiently by using a steady state Kalman Filter (KF). In a steady state KF, state covariance is computed only once during initialization. Furthermore, due to state decoupling, it is faster to linearize dynamics numerically. Details are presented in [4].

We developed an optimization based real-time walking controller for a full size 3D humanoid robot. The controller consists of two levels of optimization, a high level trajectory optimizer that reasons about center of mass and swing foot trajectories, and a low level controller that tracks those trajectories by solving a floating base full body inverse dynamics problem using Quadratic Programming. Our controller is capable of walking on rough terrain, and also achieves longer foot steps, faster walking speed, heel-strike and toe push-off. Our approach can do the VRC rough terrain walking task on our computers on a wide variety of simulated terrains. Details are presented in [1]. This paper won “Best Oral Paper Award” at Humanoids 2013.

We showed that optimal stepping trajectories and trajectory cost for a walking biped robot on rough terrain can be encoded as simple quadratic functions of initial state and footstep sequence. In order to find this encoding, we built a database of optimal walking trajectories for a 3D humanoid model by sampling the input space (initial state and footstep sequence) and solving a physically-based trajectory optimization problem for each sample. Then, the function coefficients were obtained by fitting the data using least squares. The performance of the proposed method was evaluated by comparing the function values with other optimal walking motion data generated with different footstep samples. As an application, we use a quadratic function to calculate the effort cost used in finding an optimal footstep sequence with an A* algorithm. Our study showed that a simple function can encode optimal walking effectively, which provides a fast alternative to online optimization of walking with full body dynamics. Details are presented in [3].

5.0 Conclusions

In addition to developing robot car driving in simulation, we generated publishable results on rough terrain walking. We developed a rough terrain footstep planner, a decoupled approach to state estimation, and an optimization based real-time walking controller for a full size 3D humanoid robot. We showed that optimal stepping trajectories and trajectory cost for a walking biped robot on rough terrain can be encoded as simple quadratic functions of initial state and footstep sequence.

This material is based on research sponsored by DARPA under agreement number FA8750-12-1-0307. The U. S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright notation thereon. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of DARPA or the U. S. Government

6.0 References

[1] S. Feng, X. Xinjilefu, W. Huang, and C. G. Atkeson. 3d walking based on online optimization. In

IEEE-RAS International Conference on Humanoid Robots (Humanoids), 2013.

[2] W. Huang, J. Kim, and C. G. Atkeson. Energy-based optimal step planning for humanoids. In

Robotics and Automation (ICRA), 2013 IEEE International Conference on, pages 3124–3129, 2013.

[3] J. Kim, N. S. Pollard, and C. G. Atkeson. Quadratic encoding of optimized humanoid walking. In

IEEE-RAS International Conference on Humanoid Robots (Humanoids), 2013.

[4] Xinjilefu, S. Feng, W. Huang, and C. Atkeson. Decoupled state estimation for humanoids using full-body dynamics. In IEEE International Conference on Robotics and Automation (ICRA), 2014.

7.0 List of Acronyms

COM – Center of Mass

COP – Center of Pressure

DP – Decoupled Dynamic Programming

DRC – DARPA Robotics Challenge

KF – Kalman Filter

LIPM – Linear Inverted Pendulum

M3 – Maximum Mobility and Manipulation

VRC – Virtual Robotics Challenge

ZMP – Zero Moment Point